

AUDIO OVERHANG REDUCTION FOR WIRELESS CALLS

5 Field of the Invention

The present invention relates generally to the field of wireless communications and, in particular, to reducing audio overhang in wireless communication systems.

10

Background of the Invention

Today's digital wireless communications systems packetize and
15 then buffer the voice communications of wireless calls. This buffering, of course, results in the voice communication being delayed. For example, a listener in a wireless call will not hear a speaker begin speaking for a short period of time after he or she actually begins speaking. Usually this delay is less than a second, but nonetheless, it is often noticeable and
20 sometimes annoying to the call participants.

Normal conversation has virtually no delay. When the speaker finishes speaking, a listener can immediately respond having heard everything the speaker has said. Or a listener can interrupt the speaker immediately after the speaker has finished saying something evoking a
25 comment. When substantial delay is introduced into a conversation, however, the flow, efficiency, and spontaneity of the conversation suffer. A speaker must wait for his or her last words to be heard by a listener and then after the listener begins to respond, the speaker must wait through the delay to begin hearing it. Moreover, if a listener interrupts the speaker,
30 the speaker will be at a different point in his or her conversation before beginning to hear what the listener is saying. This can result in confusion

and/or wasted time as the participants must stop speaking or ask further questions to clarify. Thus, substantial delay degrades the efficiency of conversations.

However, some delay is a necessary tradeoff in today's wireless communication systems primarily because of the error-prone wireless links. To reduce the number of voice packets that are lost, leaving gaps in the received audio, wireless systems use well-known techniques such as packet retransmission and forward error correction with interleaving across packets. Both techniques require voice packets to be buffered, and thus result in the introduction of some delay. Today's wireless system architectures themselves introduce variable delays that would distort the audio without the use of some buffering to mask these timing variations. For example, packet delivery times will vary in packet networks due to factors such as network loading. Variable delays of voice packets can also be caused by intermittent control signaling that accompanies the voice packets and as a result of a receiving MS handing off to a neighboring base site. Thus, wireless systems are designed to tradeoff the delay that results from a certain level of buffering in order to derive the benefits of providing continuous, uninterrupted voice communication.

Buffering above this optimal level, however, increases the delay experienced by users without any benefits in return. Audio buffered above this optimal level is referred to as "audio overhang." Such audio overhang can occur in wireless systems in certain situations. For example, variability in the time that some wireless systems take to establish wireless links during call setup can result in buffering with audio overhang. Because of the increased delay introduced by audio overhang, the quality of service experienced by these users can suffer substantially. Therefore, there exists a need for reducing audio overhang in wireless communication systems.

Brief Description of the Drawings

FIG. 1 is a block diagram depiction of a wireless communication system in accordance with an embodiment of the present invention.

FIG. 2 is a logic flow diagram of steps executed a wireless communication system in accordance with an embodiment of the present invention.

Description of Embodiments

To address the need for reducing audio overhang in wireless communication systems, the present invention provides for the deletion of silent frames before they are converted to audio by the listening devices. The present invention only provides for the deletion of a portion of the silent frames that make up a period of silence or low voice activity in the speaker's audio. Voice frames that make up periods of silence less than a given length of time are not deleted.

The present invention can be more fully understood with reference to FIGs. 1 and 2. FIG. 1 is a block diagram depiction of wireless communication system 100 in accordance with an embodiment of the present invention. System 100 comprises a system infrastructure, fixed network equipment (FNE) 110, and numerous mobile stations (MSs), although only MSs 101 and 102 are shown in FIG. 1's simplified system depiction. MSs 101 and 102 comprise a common set of elements. Receivers, processors, buffers (i.e., portions of memory), and speakers are all well known in the art. In particular, MS 102 comprises receiver 103, speaker 106, frame buffer 105, and processor 104 (comprising one or more memory devices and processing devices such as microprocessors and digital signal processors).

FNE 110 comprises well-known components such as base sites, base site controllers, a switch, and additional well-known infrastructure equipment not shown. To illustrate the present invention simply and concisely, FNE 110 has been depicted in block diagram form showing only receiver 111, processor 112, frame buffer 113, and transmitter 114. Virtually all wireless communication systems contain numerous receivers, transmitters, processors, and memory buffers. They are typically implemented in and across various physical components of the system. Therefore, it is understood that receiver 111, processor 112, frame buffer 113, and transmitter 114 may be implemented in and/or across different physical components of FNE 110, including physical components that are not even co-located. For example, they may be implemented across multiple base sites within FNE 110.

Operation of an embodiment of system 100 occurs substantially as follows. MSs 101 and 102 are in wireless communication with FNE 110. For purposes of illustration, MSs 101 and 102 will be assumed to be involved in a group dispatch call in which the user of MS 101 has depressed the push-to-talk (PTT) button and is speaking to the other dispatch users of the talkgroup. One of these users is the user of MS 102 who is listening to the MS 101 user speak via speaker 106. Receiver 111 receives the voice frames that convey the voice information of the call from MS 101. Some of these frames are so-called “silent frames.” In one embodiment, these frames have been marked by MS 101 to indicate that they convey either low voice activity or no voice activity. Depending on how the voice frames are voice encoded (or vocoded) these silent frames may be frames that are flagged by the vocoder as minimum rate frames (e.g., 1/8 th rate frames) or flagged as silence suppressed frames. Additionally, the silent intervals may be conveyed through the use of time stamps on the non silent frames such that the silent frames do not need to be actually sent.

Processor 112 stores the voice frames in frame buffer 113 after they are received. When frames are ready for transmission to MS 102, processor 112 extracts them and instructs the transmitter to transmit the extracted voice frames to MS 102. In similar fashion, receiver 103 then
5 receives the voice frames from FNE 110, and processor 104 stores them in frame buffer 105. The voice frames may be received by receiver 103 via Radio Link Protocol (RLP) or Forward Error Correction. As required to maintain the stream of audio for MS 102's user, processor 104 also regularly extracts the next voice frame from frame buffer 105 and de-
10 vocodes it to produce an audio signal for speaker 106 to play.

In order to reduce the audio overhang time, however, the present invention provides for the deletion of some of the silent frames before they are used to generate an audio signal. In one embodiment, the present invention is implemented in both the FNE and the receiving MS, although
15 it could alternatively be implemented in either the FNE or the MS. If implemented in both, then both processor 104 and processor 112 will be monitoring the number of voice frames stored in frame buffer 105 and frame buffer 113, respectively, as frames are being added and extracted. When the number of frames stored in either buffer exceeds a
20 predetermined size threshold (e.g., 300 milliseconds worth of voice frames), then processor 104 / 112 attempts to delete one or more silent frames.

There are a number of embodiments, all of which or some combination of which may be employed to delete silent frames. In one
25 embodiment, processor 104 / 112 scans frame buffer 105 / 113 for consecutive silent frames longer than a predetermined length (e.g., 90 msecs) and deletes a percentage (e.g., 25%) of the consecutive silent frames that exceed this length. In another embodiment, processor 104 / 112 monitors the voice frames as they are stored in the buffer. Processor
30 104 / 112 determines that a threshold number of consecutive silent frames have been stored in the frame buffer and deletes a percentage of

subsequent consecutive silent frames as they are being received and stored. In another embodiment, the deletion processing is triggered by the receipt of the last voice frame of each dispatch session within the dispatch call. Processor 104 / 112 determines that a threshold number of silent
5 frames have been consecutively stored in the frame buffer prior to the last voice frame and deletes a percentage of prior consecutive silent frames.

Regardless which deletion embodiment(s) are implemented, deleting silent frames from either frame buffer has the effect of removing that portion of the audio from what speaker 106 would otherwise play.
10 Thus, the pauses in the original audio captured by MS 101, at least those of a certain length or longer, are shortened, and audio overhang thereby reduced. While the benefits of reduced overhang are clear (as discussed in the *Background* section above), the shortening of pauses or gaps in a user's speech as received by listeners may not be desirable to some
15 users. Thus, this overhang reduction mechanism may need to be implemented as a user selected feature that can be turned on and off by mobile users.

Another ill effect of audio overhang is that in a group dispatch call, the listening users wait for the speaking user's audio, as played by their
20 MS, to complete before attempting to press the PTT to become the speaker of the next dispatch session of the call. The greater the audio overhang the longer the listener waits before trying to speak. To address this inefficiency, when MS 102 receives the last voice frame of a dispatch session within the call, MS 102 indicates to its user that the dispatch
25 session has ended and that another dispatch session may be initiated. This indication may be visual (e.g., using the display), auditory (e.g., a beep or tone), or through vibration, for example. A listener could press his or her PTT upon such an indication, the MS discard the previous speaker's unplayed audio, and the new speaker begin speaking to the
30 group without the overhang delay.

FIG. 2 is a logic flow diagram of steps executed a wireless communication system in accordance with an embodiment of the present invention. Logic flow 200 begins (202) with a communication device (an MS and/or FNE) intermittently receiving (204) and storing voice frames in a frame buffer, as it does throughout the duration of a wireless call. When (206) the audio overhang feature is enabled, the number of frames stored in the buffer is monitored (208). When (210) the number stored exceeds a threshold or maximum number, then the wireless call is developing overhang, and thus delay beyond what is optimal. To reduce this overhang, the communication device, in the most general embodiment, scans (212) the frame buffer for groups of consecutive silent frames. For the groups that are longer than a minimum silence period, a percentage of the silent frames that are in excess of the minimum silence period are deleted (214). Thus, the overhang is reduced. Throughout the wireless call, then, the communication device is monitoring for an overhang condition and deleting silent frames when an overhang condition develops.

While the present invention has been particularly shown and described with reference to particular embodiments thereof, it will be understood by those skilled in the art that various changes in form and details may be made therein without departing from the spirit and scope of the present invention.

What is claimed is:

25